

Manuel Corpas, Ph.D.

Address: ideaSpace South, University of Cambridge Biomedical Innovation Hub,
Cambridge Biomedical Campus (Clifford Allbutt Building), Hills Road
CB2 0AH Cambridge, UK.

Email: mc@manuelcorpas.com
Web: <https://manuelcorpas.com>
Mobile: +44 7939 807 507
ORCID: 0000-0002-4417-1018
Publons: 1168880
Twitter: @manuelcorpas
LinkedIn: manuelcorpas
Skype: manu_corpas

Education

2002-2007 PhD Bioinformatics, University of Manchester, UK (with internship at University of California, San Diego, USA)

2000-2002 MSc Bioinformatics, University of Manchester, UK (MSc research project at Virginia Tech, Virginia, USA)

1999-2000 Erasmus program, University of Southampton, UK

1995-1999 BSc (Hons) Biomedical Sciences, University of Navarre, Spain

Positions Held

- **Dec 2017 - Present** *Cambridge Precision Medicine (CPM), Cambridge, UK (Founder / Chief Scientist)*. Cambridge Precision Medicine helps clinicians interpret genomics for accurate diagnoses, with the objective of enabling their patients live longer and happier lives. At Cambridge Precision Medicine we manage and simplify the whole process of genetic screening, from sample to report interpretation. My work here involves the design, implementation and provision of personal genomics services, with particular emphasis on the interpretation and clinical interventions derived from the utilisation of whole genome data applied to disease prevention.
- **Jul 2018 - Present** *Institute of Continuing Education Tutor, University of Cambridge, UK*. I provide part-time and short courses for professionals on Precision Medicine and Computational Genomics. I design both face-to-face and online training materials tailored to health professionals who would like to incorporate genomics as part of their practice.
- **May 2019 - Present** *Course Director, Clinical Genetics and Personalised Medicine, Universidad Internacional de La Rioja (UNIR), Madrid, Spain*. I direct this online course in Clinical Genetics and Personalised Medicine at UNIR - the leading online university of the Spanish speaking world. Part of my role as course director involves the development of its syllabus, practical sessions, coordination with external teachers and examination of students. We utilise a virtual online platform to facilitate the dissemination of knowledge and the interactivity between

students and tutors across many countries in different timezones.

- **Sep 2016 – Dec 2017** *Repositive.io, Cambridge, UK (Scientific Lead)*. My goals as Scientific Lead included:
 1. Building the customer base around the Repositive platform
 2. Translating customer data needs into enhancements of the service we provide
 3. Understanding Precision Medicine Applications and Innovations and how to match them through proven business strategies
 4. Advocating current challenges in the implementation of preventative, predictive, participatory health solutions
- **Feb 2012 – Sep 2016** *The Genome Analysis Centre (TGAC), Norwich, UK (Project Leader)*. Next Generation Sequencing data analysis, expression and regulation mechanisms of various organisms, including human health. Development of visualisation methods for -Omics Big Data. Development of tools for analysis of personal genomes. Led a team of 2 PhD students, 2 MSc students and 2 software developers.
- **Sep 2014 – Feb 2016** *Technical Coordinator, ELIXIR-UK*. Coordination implementation of Big Data infrastructure to support the establishment of sustainable bioinformatics resources Europe-wide.
- **Apr 2011 – Jan 2015** *www.itico.org*. Director of Itico, a non-for-profit organisation registered in Cambridge, England. Development of the iAnn platform, recently adopted by ELIXIR and GOBLET (Global Organisation for Bioinformatics Learning, Education and Training).
- **Sept 2008 – Jan 2012** *Wellcome Sanger Institute, Cambridge, UK (Senior Web Developer/Postdoctoral Scientist; Lead Scientist: Dr Nigel Carter, Lead Clinician: Dr Helen Firth)*. Research and development of DECIPHER (<http://decipher.sanger.ac.uk>), a clinical genomics database for diagnosis of patients with rare genomic disorders. Developed bioinformatics tools for visualisation and analysis of personal genomes.
- **Jan-Aug 2008** *European Bioinformatics Institute, UK (Scientific Programmer; supervisors: Dr Pascal Kahlem and Dr Ewan Birney)*. Set up a systems biology database, EnDICTION, for integration of computer-generated protein-protein interactions predictions from the ENFIN network of Excellence.
- **Aug-Dec 2007** *Spanish National Bioinformatics Institute, Spain (Scientific Coordinator for the institute, led by Alfonso Valencia)*. Developed the Institute's portal and delivered strategic training of Spanish bioinformatics tools throughout the country.
- **Jul-Sep 2003** *San Diego Supercomputer Center, University of California at San Diego, USA (Internship at the Protein Data Bank in Phil Bourne's group)*. Studied the relationship of structural homology vs. sequence homology using PDB and PRINTS databases data. Funded by the World Universities Network.
- **Apr 2002 – Jun 2007** *Faculty of Life Sciences and Faculty of Computer Science, University of Manchester, UK (PhD Student; supervisors: Prof Terri Attwood and Dr Steve Pettifer)*. Study of structural factors contributing most to sequence

conservation in protein families. Created web tools that identify conserved sequence regions. Funded by the Department of Computer Science at the University of Manchester.

- **Jul-Oct 2001** *Virginia Bioinformatics Institute (Internship for MSc Dissertation; Supervisor, Prof Pedro Mendes)*. Parser in C++ for GEPASI (a biochemical pathway simulator) to allow automatic exchange of data with Systems Biology Markup Language (SBML)

Publications

Total: 56
Citations: 1998
h-index: 16

First/last/corresponding author: 31

1. **M Corpas**, S Beck, GG Glusman, M Shabani (2021) Editorial: Personal Genomes: Accessing, Sharing, and Interpretation. *Frontiers in Genetics (in production)*.
2. **M Corpas**, K Megy, V Mistry, A Metastasio, E Lehmann (2021) Whole Genome Interpretation for a Family of Five. *Frontiers in Genetics 2021; 12: 535123*.
3. V Soriano*, P Cardona, **M Corpas** (2020) A Key Action Plan for Education in a Global Crisis. *Radical Solutions for Education in a Crisis Context, 263-272*.
4. N Kovalevskaya, C Whicher, **M Corpas**, R Shaw, and F Nielsen (2018) Simplifying access to human genomic data: community platform Repositive. io. *European Journal of Human Genetics (pp. 714–714). 2018*.
5. S Beck*, AM Berner, G Bignell, M Bond, MJ Callanan, O Chervova, L Conde, **M Corpas**, S Ecker, HR Elliott, SA Fioramonti, AM Flanagan, R Gaentzsch, D Graham, D Gribbin, JA Guerra-Assunção, R Hamoudi, V Harding, PL Harrison, J Herrero, J Hofmann, E Jones, S Khan, J Kaye, P Kerr, E Libertini, L McCormack, I Moghul, N Pontikos, S Rajanayagam, K Rana, M Semega-Janneh, CP Smith, L Strom, S Umur, AP Webster, K Wint, J Wood (2018) PGP-UK: a research and citizen science hybrid project in support of personalized medicine. *BMC Med Genomics. 2018 Nov 27;11(1):108. doi: 10.1186/s12920-018-0423-1*.
6. A Bueno, R Rodríguez-López, A Reyes-Palomares, E Rojano, **M Corpas**, J Nevado, P Lapunzina, F Sánchez-Jiménez and JAG Ranea* (2018) Phenotype-loci associations in networks of patients with rare disorders: application to assist in the diagnosis of novel clinical cases. *European Journal of Human Genetics 26 (10), 1451*
7. A Thorogood*, J Bobe, B Prainsack, A Middleton, E Scott, S Nelson, **M Corpas**, N Bonhomme, L Lyman Rodriguez, M Murtagh, and Erika Kleiderman on behalf of the Participant Values Task Team of the Global Alliance for Genomics and Health (2018). Let's APPLaUD: Access for Patients and Participants to Individual-Level Uninterpreted Genomic Data. *Human Genomics, Feb 17;12(1):7*

8. **M Corpas***, NV Kovalevskaya, FGG Nielsen (2018) A FAIR Guide for Data Providers to Maximise Sharing of Human Genomic Data. *PLOS Comput Biol* 2018 Mar 15;14(3):e1005873
9. H Artaza, JM Hancock, RC Jimenez, **M Corpas*** (2017). PIsCO: A Performance indicators framework for collection of bioinformatics resource metrics. *PeerJ Preprints*
10. R Shaw, **M Corpas*** (2017). A Collection Of 2,280 Public Domain (CC0) Curated Genotypes. *bioRxiv*, 127241
11. M Tamiru, S Natsume, H Takagi, B White, H Yaegashi, M Shimizu, K Yoshida, A Uemura, Ki Oikawa, A Abe, N Urasaki, H Matsumura, P Babil, S Yamanaka, R Matsumoto, S Muranaka, G Girma, A Lopez-Montes, M Gedil, R Bhattacharjee, M Abberton, PL Kumar, I Rabbi, M Tsujimura, T Terachi, W Haerty, **M Corpas**, S Kamoun, G Kahl, H Takagi, R Asiedu, R Terauchi (2017). Genome sequencing of the staple food crop white Guinea yam enables the development of a molecular marker for sex determination. *BMC Biology* 15 (1), 86
12. RC Jiménez*, M Kuzak, M Alhamdoosh, M Barker, B Batut, M Borg, S Capella-Gutierrez, NC Hong, M Cook, **M Corpas**, M Flannery, L Garcia, JL Gelpí, S Gladman, C Goble, M González-Ferreiro, A Gonzalez-Beltran, PC Griffin, B Grüning, J Hagberg, P Holub, R Hooft, J Ison, DS Katz, B Leskošek, F López-Gómez, LJ Oliveira, D Mellor, R Mosbergen, N Mulder, Y Perez-Riverol, R Pergl, H Pichler, B Pope, F Sanz, MV Schneider, V Stodden, R Suchecki, R Svobodová-Vařeková, HA Talvik, I Todorov, A Treloar, S Tyagi, M van Gompel, D Vaughan, A Via, Xi Wang, NS Watson-Haigh, S Crouch. (2017) Four simple recommendations to encourage best practices in research software. *F1000Research*. 2017;6:ELIXIR-876.
13. C Horro, M Cook, TK Attwood, MD Brazas, JM Hancock, P Palagi, **M Corpas**, R Jimenez* (2017). BioCIDER: a Contextualisation INDEX for biological Resources discovery. *Bioinformatics*. 2017 Aug 15;33(16):2607-2608
14. MD Brazas, C Brooksbank, RC Jimenez, S Blackford, PM Palagi, J De Las Rivas, BFF Ouellette, J Kumuthini, E Korpelainen, F Lewitter, CWG van Gelder, N Mulder, **M Corpas**, MV Schneider, TW Tan, D Clements, A Davies, TK Attwood (2017). A global perspective on bioinformatics training needs. *bioRxiv*, 098996
15. Schneider MV, TK Attwood, **M Corpas**, M Charleston, S Dayalan, P Gaudet, J Kumuthini, S Morgan, PM Palagi, G Rustici. "GOBLET Standards Committee: best practices and standards in bioinformatics and biocuration". *F1000Research* 6 (2017).
16. H Artaza, N Chue Hong, **M Corpas***, A Corpuz, R Hooft, RC Jimenez, B Leskošek, BG Olivier, J Stourac, R Svobodová-Vařeková*, T Van Parys (2016). Top 10 metrics for life science software good practices. *F1000Research* 2016,5(ELIXIR):2000
17. A Reyes-Palomares, A Bueno, F Sanchez-Jimenez, MA Medina, **M Corpas**, JAG Ranea* (2016). Systematic Identification of Phenotypically Enriched Loci Using a Patient Network of Genomic Disorders. *BMC Genomics* 2016, 17:232
18. **M Corpas*** (2016) Progress on the genome characterisation of *D. alata*. *Plant*

19. **M Corpas***, W Valdivia-Granda, N Torres, B Greshake, S. Reddy, A Coletta, A Knaus, AP Harrison, M Cariaso, F Moran, F Nielsen, D Swan, DY Weiss-Solís, P Krawitz, F Schacherer, P Schols, H Yang, P Borry, G Glusman, PN Robinson (2015). Crowdsourced Analysis of Family Genomes. *BMC Genomics*
20. G Yachdav, T Goldberg, S Wilzbach, D Dao, I Shih, S Choudhary, S Crouch, M Franz, A García, LJ García, BA Grüning, D Inupakutika, I Sillitoe, AS Thanki, B Vieira, JM Villaveces, MV Schneider, S Lewis, S Pettifer, B Rost, **M Corpas*** (2015). BioJS: Anatomy of a Life Science Open Source Community. *eLife* 4, e07009
21. G Rallapalli, Fraxinus Players, DGO Saunders, K Yoshida, A Edwards, CA Lugo, S Collin, B Clavijo, **M Corpas**, D Swarbreck, M Clark, JA Downie, S Kamoun, Team Cooper, D MacLean (2015). Fraxinus: a citizen science game where players improve software DNA alignments & a model for strategising in crowdsourcing. *eLife* 4, e07460
22. F Psomopoulos*, AM Duarte*, C Blanchet, AM Bonvin, **M Corpas**, A Franc, RC Jimenez, J Marco de Lucas, T Nyrönen, G Sipos, SB Suhr (2015). Future opportunities and future trends for e-infrastructures and life sciences: going beyond grid to enable life science data analysis. *Front. Genet.* 6:197
23. TK Atwood, E Bongcam-Rudloff, ME Brazas, **M Corpas**, P Gaudet, F Lewitter*, N Mulder, PM Palagi, MV Schneider, CW van Gelder, The GOBLET Consortium (2015). GOBLET: The Global Organisation for Bioinformatics Learning, Education and Training. *PLoS Comput Biol.* 2015 Apr 9;11(4):e1004143
24. A Budd*, **M Corpas***, MD Brazas*, FC Fuller*, J Goecks, NJ Mulder, M Michaut, BF Ouellette, A Pawlik, N Blomberg (2015). A quick guide for building a successful bioinformatics community. *PLoS Comput Biol* 2015 Feb 5;11(2):e1003972
25. A Budd*, H Dinkel, **M Corpas**, FC Fuller, L Rubinat, DP Devos, PH Khoueiry, KU Förstner, F Georgatos, F Rowland, M Sharan, JX Binder, T Grace, K Traphagen, A Gristwood, NT Wood (2015). Ten simple rules for organizing an unconference. *PLoS Comput Biol* 2015 Feb 5;11(2):e1003972
26. **M Corpas***, RC Jimenez, E Bongcam-Rudloff, A Budd, MD Brazas, PL Fernandes, B Gaeta, C van Gelder, E Korpelainen, F Lewitter, A McGrath, D MacLean, PM Palagi, K Rother, J Taylor, A Via, M Watson, MV Schneider, TK Attwood (2014). The GOBLET training portal: a global repository of bioinformatics training materials, courses and trainers. *Bioinformatics*, btu601
27. D Saunders, K Yoshida, C Sambles, R Glover, B Clavijo, **M Corpas**, D Bunting, S Dong, M Clark, D Swarbreck, S Ayling, M Bashton, S Collin, T Hosoya, A Edwards, L Crossman, G Etherington, J Win, L Cano, D Studholme, JA Downie, M Caccamo, S Kamoun, D Maclean* (2014). Crowdsourced analysis of ash and ash dieback through the Open Ash Dieback project: A year 1 report on datasets and analyses contributed by a self-organising community. *BioRxiv* 2014
28. **M Corpas*** (2014). The BioJS article collection of open source components for biological data visualisation. *F1000Research* 2014, 3:56

29. **M Corpas***, RC Jimenez, SJ Carbon, A García, L Garcia, T Goldberg, J Gomez, A Kalderimis, SE Lewis, I Mulvany, A Pawlik, F Rowland, Gustavo Salazar, F Schreiber, I Sillitoe, WH Spooner, A Thanki, JM Villaveces, G Yachdav, Henning Hermjakob* (2014). BioJS: an open source standard for biological visualisation – its status in 2014. *F1000Research* 2014, **3**:55
30. AS Thanki*, S Caim, **M Corpas**, RP Davey (2014). *DNAContentViewer* a BioJS component to visualise GC/AT Content. *F1000Research* 2014, **3**:54
31. AS Thanki*, RC Jimenez, GG Kaithakottil, **M Corpas**, RP Davey (2014). *wigExplorer*, a BioJS component to visualise wig data. *F1000Research* 2014, **3**:53
32. RC Jimenez, JP Albar, J Bhak, MC Blatter, T Blicher, MD Brazas, C Brooksbank, A Budd, J De Las Rivas, J Dreyer, M van Driel, MJ Dunn, PL Fernandes, CWG van Gelder, H Hermjakob, V Ioannidis, DP Judge, P Kahlem, E Korpelainen, HJ Kraus, J Loveland, C Mayer, J McDowall, F Moran, N Mulder, T Nyronen, K Rother, GA Salazar, MV Schneider, R Schneider, A Via, JM Villaveces, P Yu, TK Attwood, **M Corpas*** (2013). iAnn: An Event Sharing Platform for the Life Sciences. *Bioinformatics* 29(15), 1919-1921
33. **M Corpas***, M Cariaso, A Coletta, D Weiss, AP Harrison, F Moran, H Yang (2013). A Complete Public Domain Family Genomics Dataset. *BioRxiv* 2013
34. J de Ridder, Y Bromberg , M Michaut, VP Satagopam, **M Corpas**, G Macintyre, T Alexandrov* (2013). The young PI buzz: learning from the organizers of the Junior Principal Investigator meeting at ISMB-ECCB 2013. *PLoS Comput Biol.* 2013;9(11):e1003350
35. **M Corpas*** (2013) Crowdsourcing the Corpasome. *Source Code for Biology and Medicine* 2013, **8**:13
36. RC Jimenez*, **M Corpas** (2013). Bioinformatics Workflows and Web Services in Systems Biology Made Easy for Experimentalists. *Methods Mol Biol.* 2013;1021:299-310
37. J Gómez, LJ García, GA Salazar, JV, S Gore, A García, MJ Martín, G Launay, R Alcántara, N Del Toro Ayllón, M Dumousseau, S Orchard, S Velankar, H Hermjakob, C Zong, P Ping, **M Corpas**, RC Jiménez* (2013). BioJS: An Open Source JavaScript Framework for Biological Data Visualization. *Bioinformatics*, *btt100*
38. **M Corpas*** (2012). A Genome Blogger Manifesto. *GigaScience* 1:15 doi:10.1186/2047-217X-1-15
39. GJ Swaminathan, E Bragin, EA Chatzimichali, **M Corpas**, AP Bevan, CF Wright, NP Carter, ME Hurles, HV Firth* (2012). DECIPHER: web-based, community resource for clinical interpretation of rare variants in developmental disorders. *Hum Mol Genet.* 2012 Oct 15;21(R1):R37-44.
40. **M Corpas***, S Fatumo, R Schneider (2012). How not to be a Bioinformatician. *Source Code for Biology and Medicine* 7:3
41. **M Corpas***, E Bragin, S Clayton, P Bevan, HV Firth (2012). Interpretation of Genomic Copy Number Variants Using DECIPHER. *Current Protocols in Human Genetics.* 72:8.14.1–8.14.17

42. **M Corpas*** (2012). Experiences with Personal Genomics: A Family Journey. *Journal of Genetic Counseling* 21(3), 386-391
43. **M Corpas*** (2012). Spanish Cuts: More Economic Damage. *Nature*. 2012 Jul 4;487(7405):38
44. G Glusman, M Cariaso, RC Jimenez, D Swan, B Greshake, J Bhak, DW Logan, **M Corpas*** (2012). Low budget analysis of Direct-To-Consumer genomic testing familial data. *F1000 Research*, 1:3 doi: 10.3410/f1000research.1-3.v1
45. RC Jimenez, GA Salazar, B Gel, P Bevan, J Dopazo, N Mulder, **M Corpas*** (2011). myKaryoView: A Light-Weight Client for Visualization of Genomic Data. *PLOS ONE* 6(10): e26345
46. **M Corpas**, S Doumbia, O Gascuel*, N Mulder (2011). Proceedings of the Bamako 2009 Conference on the Bioinformatics of Infectious Diseases. *Infection, Genetics and Evolution. Elsevier* 11(4), 695-697
47. HV Firth*, D R Fitzpatrick, M Parker, AP Bevan, **M Corpas**, E Bragin, S Clayton, J Barrett, ME Hurles, and NP Carter* (2010). The Deciphering Developmental Disorders (DDD) project. *J Med Genet*, vol. 47, pp. S43-S43
48. HV Firth*, SM Richards, AP Bevan, S Clayton, **M Corpas**, D Rajan, S Van Vooren, Y Moreau, RM Pettett, NP Carter* (2009). DECIPHER: DatabasE of Chromosomal Imbalance and Phenotype using Ensembl Resources. *Am J Hum Genet*. 2009 Apr;84(4):524-33
49. **M Corpas**, S Richards, AP Bevan, S Van Vooren, RM Pettett, HV Firth, and NP Carter* (2009). DECIPHER: Shedding Light on Chromosomal Imbalance and Phenotype Interpretation. *Chromosome Research*, vol. 17, pp. 13-14
50. **M Corpas**, N Gehlenborg, SC Janga, PE Bourne* (2008). Ten Simple Rules for Organizing a Scientific Meeting. *PLOS Comput Biol* 4(6): e1000080
51. F Reisinger, **M Corpas**, J Hancock, H Hermjakob, E Birney, P Kahlem* (2008). ENFIN – an Integrative Structure for Systems Biology. *Lecture Notes in Computer Science (DILS), Volume 5109/2008* 132-143
52. N Gehlenborg*, **M Corpas**, SC Janga (2007) – Highlights from the Third International Society For Computational Biology Student Council Symposium at the Fifteenth Annual International Conference on Intelligent Systems for Molecular Biology. *BMC Bioinformatics* 2007, 8(Suppl 8):11
53. **M Corpas**, J Sinnott, D Thorne, S Pettifer, and TK Attwood*, and the PFF Consortium (2007). PFF – An integrated resource of residues and fragments critical for protein folding. *BMC Systems Biology* 2007, 1(Suppl 1):P48
54. **M Corpas** (2007). Folding Patterns in Protein Sequences. *PhD thesis submitted to the Computer Science Department of the University of Manchester, Manchester, UK.*
55. **M Corpas*** (2005). Community Outreach. *Nature* 7054-1204b, 24 August 05
56. **M Corpas*** (2003). Integrating Simulation Packages via Systems Biology Markup Language. *Lecture Notes in Computer Science, 2003, Volume 2602, Computational Methods in Systems Biology*

*corresponding author

Book Chapter

- R.C. Jimenez, **M. Corpas** (2013). Bioinformatics Workflows and Web-services in Systems Biology Made Easy. *In Silico Systems Biology. Methods in Molecular Biology, Volume: 1021, DOI: 10.1007/978-1-62703-450-0_16*

Competitive Grants Awarded

- 2020** European Commission (EASI-Genomics): whole exome sequencing of 100 COVID-19 patients from the first wave of the infection, with the possibility of further funds for whole genome sequencing of up to 300 samples. Estimated value: euro 100K. Role co-PI, lead the genome analysis part.
- 2017** NIH Data Commons Pilot Phase: FAIR Data to Drive Cures (total grant value; \$800K). PI: Brandi Davis-Dusenbery (Seven Bridges). Role in the grant: M Corpas led the writing of two key capabilities (i.e., work packages) where Repositiv leads.
- 2016** Google Summer of Code to develop BioJavaScript functionality (value: \$33K USD)
- 2015** Horizon 2020, Infradev-3, EXCELERATE (euro 180K). Leader of The Genome Analysis Centre's contribution to the ELIXIR-EXCELERATE programme
- 2015** Norwich Research Park Translational Fund (£50K). Partnership with Repositiv.io to develop interfaces for visualisation of distributed clinical genomics data
- 2014** BBSRC ICASE PhD fellowship funding, 4 years (value £92K) "Living Figures". Industrial Partner: eLife journal. eLife to support stipend fees for student (£56)
- 2014** Google Summer of Code, 5 internships to develop BioJavaScript functionality (value \$33K USD)
- 2013** BBSRC ICASE PhD fellowship funding, 4 years (value £92K) "Integrating Visualisation of Unreferenced Genomes". Industrial Partner: Eagle Genomics
- 2013** Software Sustainability Institute Consultancy Call to support the development of the BioJavaScript project (3 person months)

Line Management

- **Nadia Kovaleskaya** *Repositiv Scientist*. Development of training materials in human genomic data sharing. Building of the community of users for the Repositiv platform.
- **Haydee Artaza** *ELIXIR Scientific Programmer*. Development of ELIXIR metrics registry to measure impact of European Core Resources. Collaboration ELIXIR-Hub/TGAC.
- **Carlos Horro** *PostDoc BioJS senior developer*. Development of BioJS interfaces for ELIXIR-EXCELERATE WP5 (Interoperability) and WP7 (Plant Genome-Phenome). Supports development of GOBLET Training Portal and ELIXIR metrics

registry

- **Jessica Jordan** *BioJS software developer*. Development of a BioJS interface for Repositio.io to allow interactive visualisation of personal genomics data.

Research Students Supervised

- **Manuela Boluda** (May-Aug 2016) MSc Thesis in Bioinformatics.
- **Dennis Swartz** (Sep 2015 – Mar 2016) *MSc Thesis: “Development of BioJS component for multilayer network visualisation”*. Collaboration with RostLab, Munich and Pietro Lio, University of Cambridge
- **Benjamin White** (Oct 2014-Feb 2016). Development of visualisation methods for analysis of yam DNA-seq and RNA-seq data. *BBSRC funded iCASE PhD Thesis, TGAC/University of East Anglia. Collaboration with Eagle Genomics*
- **Evanthia Kaimaklioti** (Oct 2015-2018). Living Figures: Development of Interactive Components for Reproducible Results in Published Articles. *BBSRC funded iCASE PhD Thesis, TGAC/University of East Anglia. Collaboration with eLife and University of Manchester*
- **Benjamin White** (Oct 2014-2017). Development of visualisation methods for analysis of yam DNA-seq and RNA-seq data. *BBSRC funded iCASE PhD Thesis, TGAC/University of East Anglia. Collaboration with Eagle Genomics*
- **Benedikt Rauscher** (May 2015 – Oct 2015) *MSc Thesis: “VAPoR: Visual Annotation of Plant Resistance”*. Collaboration with RostLab, Munich
- **Aníbal Bueno** Visiting PhD student (University of Málaga). Development of BioJS interfaces for clinical genomics analyses
- **Thanos Mourakis** (2013). Development of pipelines for NGS analysis of Microbes. *TGAC/MSc Thesis, University of East Anglia.*
- **Jesús Torres García** (2011). Collection of publicly available 23andMe personal genotypes and parsing of features. *Undergraduate internship, Sanger Institute*

Community Outreach/Academic Service

- 2021** Organiser of Wellcome Genome Conference on Personal Genomics, Hinxton, UK
- 2020** Editor of Research Topic Article Collection on Personal Genomics, *Frontiers in Genetics Journal*, Lausanne, Switzerland
- 2019** Scientific Committee, Longevity World Forum, Valencia, Spain
- 2018** Scientific Committee, Longevity World Forum, Valencia, Spain
- 2018** Member of the British Society for Genetic Medicine (BSGM), London, UK
- 2018** Organiser of Wellcome Genome Conference on Personal Genomics, Hinxton, UK
- 2017** Member of ELIXIR-UK SME Programme, Cambridge, UK
- 2017** Member of the Global Alliance for Genomics and Health (GA4GH), Ethics and Regulatory Committee
- 2016** Fellow of the Software Sustainability Institute, UK
- 2016** Faculty 1000 Advisory Board Member in Computational Genomics and Genetic Analysis
- 2015** Steering Committee of BiVi, the UK Biological Visualisation Network

- 2015** Chair BioJS Annual Conference (Norwich, UK)
- 2015** Scientific Committee, Galaxy Community Conference (Norwich, UK)
- 2015** Chair ISCB Africa ASBCB Conference on Bioinformatics (Dar-es-Salaam, Tanzania)
- 2014** Board of Directors, International Society for Computational Biology (ISCB)
- 2013** Technical Committee Chair for the Global Organisation for Bioinformatics Learning, Education and Training (GOBLET)
- 2013** Organiser of Networking Workshop at ISMB* conference, Berlin, Germany
- 2013** Junior PIs workshop organizer, ISMB conference 2013, Berlin, Germany
- 2013** Steering Committee ISCB* Africa ASBCB Conference in Bioinformatics, Tunis, Tunisia
- 2011** Contributor to the Computational Biology WikiProject, Wikipedia
- 2011** Steering Committee ISCB Africa ASBCB Conference in Bioinformatics, Cape Town, South Africa
- 2010** Advisor Joint Regional Student Groups Committee for ISCB Africa
- 2010** Organizing Committee First North Africa, Portugal and Spain Student Symposium on Bioinformatics
- 2009** co-Chair ISCB-Africa ASBCB Joint Conference on Bioinformatics of Infectious Diseases, Bamako, Mali, West Africa
- 2008** Organizing Committee Visual Reflections on Science, ISMB, Toronto, Canada
- 2007** Chair of the Computational Biology Student Council Symposium 3, Vienna, Austria
- 2007** Local Organising Committee of BioSysBio Conference, Manchester, UK
- 2006, 2008** Representative to the Board of Directors of the ISCB Student Council
- 2004** Inaugural Chair for the ISCB Student Council, Glasgow, UK

*ISCB: International Society for Computational Biology

*ASBCB: African Society for Bioinformatics and Computational Biology

*ISMB: Intelligent Systems for Molecular Biology

Journal Editorial Board

- PLOS ONE
- BMC Medical Genomics
- F1000Research
- Frontiers in Genetics

Fellowships/Honours/Awards

Software Sustainability Institute Fellowship Award (17 awardees out of 133 applicants), 2016

Award: £3K

Performance Award for the year, The Genome Analysis Centre (2015)

Performance Award for Contribution to DECIPHER Project, Sanger Institute (2011)

Top 15 Bioinformatics Blog in 2010, Manuel Corpas' Blog (voted by NovoSeek)

WTSI Performance Award for Contribution to DECIPHER Project, Sanger Institute (2010)

Internship, World Universities Network, Protein Data Bank, San Diego Supercomputer Center, University of California, San Diego (2003)

Founder ISCB Student Council (2004)

PhD Studentship, Computer Science Dept. University of Manchester, UK (2002)

Travel Grant, Protein Folding Fragment Project, QLG2-CT-2002-1298. (2002)

Socrates-Erasmus. Exchange year, University of Southampton, UK (1999)